

# Generalized Riordan Groups and Operators on Polynomials

Shaul Zemel\*

May 13, 2015

## Introduction

Sequences of polynomials appear in many branches of mathematics and physics, e.g., the ones named after Bernoulli, Euler, Hermite, Lagrange, Laguerre, and others. Many of them are known to be *Sheffer sequences*, one of its equivalent defining properties is that the linear operator taking  $p_n$  to  $np_{n-1}$  commutes with the translation operators. The umbral calculus was invented, initially symbolically and afterwards more rigorously, in order to understand the combinatorics of such sequences of polynomials, and prove identities between them and between their coefficients. One feature of these sequences is that their coefficients may be seen as the entries of certain infinite lower triangular matrices called *Riordan arrays*.

There is a vast literature on the subject, large parts of which are aimed at applications to special sequences of polynomials. We first mention [R], which provides a clear introduction to the basics of this theory. A very partial list of references, in which one may find ways to deduce interesting results from the Riordan (or Sheffer) property consists of [LM], [S], and the references cited there. In these references the Sheffer sequences are related to some shift operators, Riordan arrays, recurrence relations, linear functionals on polynomials, and further algebraic and combinatorial objects.

In this paper, however, we are interested in a more algebraic approach to this theory. [SGWW] considers the set of Riordan arrays as a group (coining the term as well). [WW] considers the more general groups of Riordan arrays, which are related to the non-classical umbral calculi appearing also in Chapter 6 of [R] (both references, and the papers cited therein, should be added to the list from the previous paragraph). We consider the set of all graded sequences of polynomials (with no additional properties) as a group of infinite, lower-triangular matrices. We then present our point of view on Sheffer sequences and Riordan arrays in the generalized sense of [WW], in particular as a subgroup of this bigger group. Next, we relate several equivalent properties that these

---

\*This work was partially supported by DFG grant BR-2163/4-1.

sequences have: The column vectors of their matrices give, using the appropriate weights, a geometric sequence of formal power series; Their weighted generating function has a succinct expression as a function of two variables; The associated derivative-like operator commutes with appropriate translation-like operators; They respect a certain product rule for linear functionals on polynomials; And their weighted dual basis takes the form of a geometric sequence.

The group of  $W$ -Sheffer sequences contains two natural subgroups: The group of  $W$ -binomial sequences (which are characterized by their values at 0), and the group of  $W$ -Appell sequences. Example of those, together with some of their properties, are also given in the references mentioned above. As abstract groups these are isomorphic to formal power series of valuation 1 with composition and formal power series of valuation 0 with multiplication respectively. The  $W$ -Sheffer sequences are their semi-direct product. This structure is the same for every weight. Now, our large group of all graded sequences admits acts on a certain set of operators. We prove that the  $W$ -Appell sequences form the stabilizer of one such operator, and the  $W$ -Riordan group is its normalizer in the larger group. As this action is transitive, this explains why the algebraic structure must indeed be independent of the weight. Note that unlike many authors, we do not assume (except in one example) that our base field is of characteristic 0.

This paper is divided to 5 sections. In Section 1 we present the large group, generalized Riordan arrays, and the relation to special forms of power series in two variables. Section 2 defines Sheffer sequences (and their subgroups) via translation and derivation operators, and proves the equivalence to Riordan arrays. In Section 3 we investigate linear functionals on polynomials, and prove the relation between Sheffer sequences and products on such functionals. Section 4 introduces the action of sequences on operators, and shows that Sheffer sequences normalize a stabilizer in this action. Finally, Section 5 gives the classical examples in this language, and also proves a result about the intersection of generalized Riordan groups with two different weights.

I am grateful to A.-M. Luzón-Cordero for her willingness to read a preliminary version of this paper, as well as for her encouragement to send it to publication.

## 1 Riordan Groups

Let  $\mathbb{F}$  be a field.  $\mathbb{F}[x]$  denotes the ring of polynomials in one indeterminate  $x$ , and  $\mathbb{F}[[y]]$  is the ring of formal power series in one indeterminate  $y$  over  $\mathbb{F}$ . The latter ring is a complete discrete valuation ring. The valuation  $v(C)$  of a non-zero element  $C(y) = \sum_{n=0}^{\infty} c_n y^n \in \mathbb{F}[[y]]$  is the minimal index  $n$  such that  $c_n \neq 0$ . An element  $C \in \mathbb{F}[[y]]$  is invertible (i.e., it has a multiplicative inverse) if and only if  $v(C) = 0$ , i.e., if and only if  $c_0 \neq 0$ .

Let two sequences  $\{p_n(x)\}_{n \in \mathbb{N}}$  and  $\{q_n(x)\}_{n \in \mathbb{N}}$  of polynomials in  $\mathbb{F}[x]$  be

given. Their *umbral composition*  $\{r_n(x)\}_{n \in \mathbb{N}}$  is defined as follows: If

$$p_n(x) = \sum_{k=0}^{d_n} a_{n,k} x^k \quad \text{with } a_{n,k} \in \mathbb{F} \quad \text{then} \quad r_n(x) = \sum_{k=0}^{d_n} a_{n,k} q_k(x)$$

(where  $d_n$  is the degree of  $p_n$ ). We consider these sequences as infinite column vectors (with integer indices  $\geq 0$ ) over  $\mathbb{F}[x]$ , where the natural “base point” for these sequences is the sequence of monomials  $m_n(x) = x^n$ . We now have

**Proposition 1.1.** *For any sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  there exists a unique infinite matrix  $A$ , in which each row has only finitely many non-zero entries, taking the monomial sequence to  $\{p_n(x)\}_{n \in \mathbb{N}}$ . Conversely, every row-finite matrix arises in this way from a unique sequence of polynomials. The umbral composition of sequences corresponds to the matrix product.*

The proof is simple and straightforward. Note that the essential finiteness of all the rows makes the matrix product in Proposition 1.1 well-defined. In addition, this property is preserved in products: If  $A$  and  $B$  correspond to the sequences  $\{p_n(x)\}_{n \in \mathbb{N}}$  and  $\{q_n(x)\}_{n \in \mathbb{N}}$  respectively, and if  $d_n$  and  $e_n$  are the degrees of  $p_n$  and  $q_n$  respectively, then  $a_{n,k} = 0$  for all  $k > d_n$  and  $b_{k,l} = 0$  for all  $l > e_k$ . Hence  $(AB)_{n,l} = 0$  wherever  $l > \max\{e_k\}_{k=0}^{d_n}$ , so that the  $n$ th row of  $AB$  also contains just finitely many non-zero entries. This corresponds to the fact that the umbral composition of two sequences of polynomials is indeed a sequence of polynomials.

We shall restrict attention to sequences of polynomials which are *graded*, i.e., in which  $d_n = n$  for all  $n$ . For these sequences we have

**Proposition 1.2.** *The sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is graded if and only if the corresponding matrix  $A$  is lower triangular and invertible. The set of such matrices forms a group  $L = L_\infty(\mathbb{F})$  under matrix multiplication.*

The proof is again simple and direct. It follows immediately from Propositions 1.1 and 1.2 that the umbral composition of two graded sequences is also graded. All our sequences henceforth will be assumed to be graded, unless explicitly stated otherwise.

Let  $W = W(t) = \sum_{n=0}^{\infty} \frac{t^n}{w_n}$  be an element of  $\mathbb{F}[[t]]$  in which  $w_0 = 1$  and  $w_n \neq 0$  for every  $n$ . Fixing such  $W$  we define, for every  $A \in L$ , the power series  $C_{A_k, W}(y) = \sum_{n=0}^{\infty} a_{n,k} \frac{y^n}{w_n} \in \mathbb{F}[[y]]$ . We shall write just  $C_{A_k}$  instead of  $C_{A_k, W}$  in case  $W$  is clear from the context. If the choice of  $A$  is also clear we may shorten  $C_{A_k}$  further to just  $C_k$ . It is clear that  $v(C_{A_k}) = k$  for each  $k$ . We now make the following

**Definition 1.3.** *We say that  $A$  is a Riordan array of weight  $W$ , or a  $W$ -Riordan array, if the equality  $w_k^2 C_{A_k, W}^2 = w_{k-1} C_{A_{k-1}, W} w_{k+1} C_{A_{k+1}, W}$  holds in  $\mathbb{F}[[y]]$  for every  $k \geq 1$ .*

We denote the set of  $W$ -Riordan arrays by  $R_W$ . We may say just a *Riordan array* in case the choice of  $W$  is clear.

Definition 1.3 is equivalent to the statement that there exist two power series  $\alpha$  and  $\beta$  in  $\mathbb{F}[[y]]$ , with  $v(\alpha) = 0$  and  $v(\beta) = 1$ , such that  $C_{A_k}$  equals  $\alpha \cdot \frac{\beta^k}{w_k}$  for all  $k$ . Indeed,  $\alpha$  is just  $C_0$ , while  $\beta$  is the quotient  $\frac{w_{k+1}C_{k+1}}{w_k C_k}$ , which is assumed to be independent of  $k$ . It is now clear that  $W$ -Riordan arrays are in one-to-one correspondence with the set of pairs of elements  $(\alpha, \beta)$  from  $\mathbb{F}[[y]]$  with  $v(\alpha) = 0$  and  $v(\beta) = 1$ .

Recall that the set  $\mathbb{F}[[y]]^\times$  of elements of valuation 0 in  $\mathbb{F}[[y]]$  forms a (commutative) group under multiplication of power series, with identity 1. In addition, the set  $y\mathbb{F}[[y]]^\times$  of elements of valuation 1 form a (non-commutative) group, with identity  $e(y) = y$ , under composition of power series. It turns out more convenient to use the convention of opposite composition, in which the product of the power series  $\beta$  and  $\delta$  is  $\delta \circ \beta$ . This composition law extends to the case where  $v(\delta) = 0$ , yielding an action of the latter group on the former.

We can now prove

**Proposition 1.4.** *The Riordan group  $R_W$  is a subgroup of  $L$ . Moreover, it is isomorphic to the semi-direct product in which  $y\mathbb{F}[[y]]^\times$  operates on  $\mathbb{F}[[y]]^\times$  as above.*

*Proof.* Consider an element  $A \in L$  and the infinite row vector whose  $n$ th entry is  $\frac{y^n}{w_n}$ . Their product is formally well-defined in  $\mathbb{F}[[y]]$ , giving the row vector with  $k$ th entry  $C_{A_k}(y)$  (by definition). Moreover, as  $v(C_{A_k}) = k$  we may multiply this matrix by another element of  $L$ , and the associative law holds for these products. Assume now that  $A$  lie in  $R_W$ , and corresponds to the pair  $(\alpha, \beta)$  of  $\mathbb{F}[[y]]^\times \times y\mathbb{F}[[y]]^\times$ . This means that the product of our row vector with  $A$  yields a row vector whose  $k$ th entry is  $\alpha(y) \frac{\beta(y)^k}{w_k}$ . But then we can take the scalar  $\alpha(y)$  out, and the same argument shows that taking the product with  $B$  yields the row vector whose  $l$ th entry is  $\alpha(y) \cdot C_{B_l}(\beta(y))$ . Assuming further that  $B \in R_W$  and that  $(\gamma, \delta)$  is the corresponding pair, we find that the latter  $l$ th entry is  $\alpha(y) \gamma(\beta(y)) \frac{\delta(\beta(y))}{w_l}$ . This shows (by associativity) that  $AB$  is the element of  $R_W$  which corresponds to the pair  $(\alpha \cdot (\gamma \circ \beta), \delta \circ \beta)$ . As this is precisely the product rule in the semi-direct product in question, this proves both assertions of the proposition.  $\square$

Taking a matrix  $A \in L$ , multiplying it from the right by our vector of monomials and from the left by a row vector of the form appearing in Proposition 1.4 yields an element of the ring  $\mathbb{F}[x][[y]]$  (this ring is larger than  $\mathbb{F}[[y]][x]$ , as the latter ring contains only expressions whose total degree in  $x$  is bounded, but both are contained in the ring  $\mathbb{F}[[x, y]]$  of formal power series in two variables). Explicitly, the result can be written either as  $\sum_n p_n(x) \frac{y^n}{w_n}$ , or, considered as an element of  $\mathbb{F}[[x, y]]$  and organized in a different order, as  $\sum_k C_{A_k, W}(y) x^k$ . Elements of the  $W$ -Riordan group can now be characterized according to the following

**Proposition 1.5.** *The matrix  $A$  lies in  $R_W$  if and only if the corresponding element in  $\mathbb{F}[x][[y]]$  is of the form  $\alpha(y)W(x\beta(y))$  for power series  $\alpha$  and  $\beta$  in  $\mathbb{F}[[y]]$  with  $v(\alpha) = 0$  and  $v(\beta) = 1$ .*

Note that as  $v(\beta) = 1$ , the expression  $\alpha(y)W(x\beta(y))$  is well-defined in  $\mathbb{F}[[x, y]]$ .

*Proof.* The element of  $\mathbb{F}[[x, y]]$  which corresponds to  $A$  is  $\sum_k C_{A_k, W}(y)x^k$ , while expanding  $\alpha(y)W(x\beta(y))$  in that ring according to powers of  $x$  yields the sum  $\sum_k \alpha(y) \frac{\beta(y)^k x^k}{w_k}$ . It is now clear that these power series coincide if and only if  $A$  is the element of  $R_W$  corresponding to the pair  $(\alpha, \beta)$ . This proves the proposition.  $\square$

## 2 Sheffer Sequences

Let  $D_W : \mathbb{F}[x] \rightarrow \mathbb{F}[x]$  be the “weighted derivative”, i.e., the linear operator which takes  $\frac{x^n}{w_n}$  to  $\frac{x^{n-1}}{w_{n-1}}$  for each  $n \geq 0$  (where  $\frac{1}{w_{-1}}$  is defined to be 0). More generally, let  $\{p_n(x)\}_{n \in \mathbb{N}}$  be a (graded) sequence of polynomials. There exists a unique operator, which we denote  $Q_W$  (or  $Q$  if  $W$  is clear from the context), sending  $\frac{p_n}{w_n}$  to  $\frac{p_{n-1}}{w_{n-1}}$  for every  $n$ . Given  $h \in \mathbb{F}$  we define the  $W$ -translation in  $h$  to be the element  $T_{h, W}$  of  $L$  taking  $\frac{x^n}{w_n}$  to  $\sum_{k=0}^n \frac{h^{n-k}}{w_{n-k}} \cdot \frac{x^k}{w_k}$ . We now make the following

**Definition 2.1.** *The sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is called a Sheffer sequence of weight  $W$ , or a  $W$ -Sheffer sequence, if the operator  $Q_W$  commutes with all the  $W$ -translations.*

If  $A \in L$  is the matrix corresponding to the sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  as in Proposition 1.1, then we may use the notation  $Q_{A, W}$  or  $Q_A$  for the appropriate operator in case confusion may arise as to the sequence to which it relates.

Given a sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$ , it is clear from the definition of  $T_{h, W}$  that the coefficient of  $x^k$  in  $T_{h, W}(p_n)$  is a polynomial in  $h$  of degree  $n - k$ . The same assertion holds in case we express  $T_{h, W}(p_n)$  using the basis  $\{p_k\}_{k \in \mathbb{N}}$  of  $\mathbb{F}[x]$ . We may therefore write

$$T_{h, W}\left(\frac{p_n}{w_n}\right)(x) = \sum_{k=0}^n \frac{d_{n, k, A, W}(h)}{w_{n-k}} \cdot \frac{p_k(x)}{w_k}, \quad \text{with } d_{n, k, A, W}(t) \in \mathbb{F}[t],$$

and the degree of  $d_{n, k, A, W}$  is  $n - k$ . Our next assertion is

**Lemma 2.2.** *The sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is  $W$ -Sheffer if and only if the polynomials  $d_{n, k, A, W}$  depend only on the difference  $n - k$ .*

*Proof.* We compare the two compositions  $T_{h, W}Q_{A, W}$  and  $Q_{A, W}T_{h, W}$  by evaluating their actions on the polynomial  $\frac{p_n}{w_n}$ . The first composition gives

$$T_{h, W}Q_{A, W}\left(\frac{p_n}{w_n}\right)(x) = T_{h, W}\left(\frac{p_{n-1}}{w_{n-1}}\right)(x) = \sum_{k=0}^{n-1} \frac{d_{n-1, k, A, W}(h)}{w_{n-1-k}} \cdot \frac{p_k(x)}{w_k},$$

while the other one yields

$$\sum_{l=0}^n \frac{d_{n,l,A,W}(h)}{w_{n-l}} Q_{A,W} \left( \frac{p_l}{w_l} \right) (x) \stackrel{l=k+1}{=} \sum_{k=0}^{n-1} \frac{d_{n,k+1,A,W}(h)}{w_{n-k-1}} \cdot \frac{p_k(x)}{w_k}$$

(where we may omit the term with  $l = 0$  and  $k = -1$  since  $Q_{A,W}$  annihilates it). Since  $\left\{ \frac{p_k(x)}{w_k} \right\}_{k \in \mathbb{N}}$  is a basis for  $\mathbb{F}[x]$ , the two sides coincide if and only if the equality  $d_{n-1,k,A,W} = d_{n,k+1,A,W}$  holds for every  $k < n$ . This proves the lemma.  $\square$

Given a sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$ , we consider the element  $\sum_n p_n(x) \frac{y^n}{w_n}$  of the ring  $\mathbb{F}[x][[y]] \subseteq \mathbb{F}[[x, y]]$  as above. From Lemma 2.2 we deduce

**Proposition 2.3.** *The sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is  $W$ -Sheffer if and only if the operator  $T_{h,W}$  multiplies the corresponding element of  $\mathbb{F}[x][[y]]$  by a power series in  $y$  (and  $h$ ), independently of  $x$ , for every  $h \in \mathbb{F}$ .*

*Proof.* If  $A$  is the element of  $L$  corresponding to  $\{p_n(x)\}_{n \in \mathbb{N}}$ , then the formula from above shows that  $T_{h,W} \left( \sum_n p_n \frac{y^n}{w_n} \right) (x)$  equals

$$\sum_n \sum_{k=0}^n y^n \frac{d_{n,k,A,W}(h)}{w_{n-k}} \cdot \frac{p_k(x)}{w_k} \stackrel{n=k+l}{=} \sum_k \sum_l \frac{d_{k+l,k,A,W}(h) y^l}{w_l} \cdot \frac{p_k(x) y^k}{w_k}.$$

By Lemma 2.2 it suffices to show that the latter expression multiplies the power series  $\sum_k p_k(x) \frac{y^k}{w_k}$  by a power series in  $y$  and  $h$  if and only if the polynomials  $d_{n,k,A,W}$  depend only on the difference  $n - k$ . Now, if  $d_{n,k,A,W} = d_{n-k}$  for some polynomial  $d_{n-k}$  (of degree  $n - k$ ) then the latter formula is the product of  $\sum_k p_k(x) \frac{y^k}{w_k}$  and the series  $\sum_l d_l(h) \frac{y^l}{w_l} \in \mathbb{F}[h][[y]]$ , as required. On the other hand, we recall that  $\frac{p_k(x)}{w_k} y^m$ , with  $k$  and  $m$  from  $\mathbb{N}$ , are linearly independent in  $\mathbb{F}[[x, y]]$ . Therefore, if  $T_{h,W}$  multiplies our expression by some element of  $\mathbb{F}[[h, y]]$ , which we write as  $\sum_l \frac{d_l(h) y^l}{w_l}$ , then the comparison of the coefficient of  $\frac{p_k(x)}{w_k} y^{k+l}$  in this product with the one appearing in our formula for  $T_{h,W} \left( \sum_n p_n \frac{y^n}{w_n} \right) (x)$  yields the equality  $d_{k+l,k,A,W} = d_l$  (of power series in  $h$ ) for all  $k$  and  $l$ . As this shows that  $d_{n,k,A,W}$  depends only on the difference  $n - k$  (and is a polynomial of degree  $n - k$ ), this completes the proof of the proposition.  $\square$

We can now relate Definitions 1.3 and 2.1 in

**Theorem 2.4.** *Let a weight  $W = \sum_n w_n t^n$  and a sequence of polynomials  $\{p_n(x)\}_{n \in \mathbb{N}}$ , with the associated matrix  $A \in L$ , be given. Then the sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is  $W$ -Sheffer if and only if the matrix  $A$  lies in  $R_W$ .*

*Proof.* By Proposition 2.3 it suffices to show that  $A \in R_W$  if and only if the operation of all translation operators  $T_{h,W}$  on the expression from Proposition 1.5 multiplies it by a power series in  $y$  (and  $h$ ) which is independent of  $x$ . Now,

if  $A \in R_W$  then Proposition 1.5 shows that the expression in question equals  $\alpha(y)W(x\beta(y))$ . Expanding the series  $W$  and letting  $T_{h,W}$  operate, we obtain

$$\alpha(y) \sum_n \beta(y)^n \sum_{k=0}^n \frac{h^{n-k} x^k}{w_{n-k} w_k} \stackrel{n=k+l}{=} \alpha(y) \sum_l \frac{h^l \beta(y)^l}{w_l} \sum_k \frac{x^k \beta(y)^k}{w_k},$$

which is the original expression multiplied by  $W(h\beta(y))$ . Conversely, recall that this power series can also be written as  $\sum_l C_{A_l, W}(y) x^l$ . Hence the operation of  $T_{h,W}$  sends it to

$$\sum_l C_{A_l}(y) \sum_{k=0}^l \frac{w_l h^{l-k} x^k}{w_{l-k} w_k} = \sum_k C_{A_k}(y) x^k + \frac{h}{w_1} \sum_k C_{A_{k+1}}(y) \frac{w_{k+1} x^k}{w_k} + O(h^2),$$

where the  $O(h^2)$  means a power series in  $x$ ,  $y$ , and  $h$  which contains only powers of  $h$  which are at least 2. Assume that the latter expression equals the product of  $\sum_k C_{A_k}(y) x^k$  with an element of  $\mathbb{F}[[h, y]]$ , and we write this multiplier as  $\gamma(y) + \frac{h}{w_1} \beta(y) + O(h^2)$ . Comparing the resulting expressions we obtain  $\gamma(y) = 1$ , as well as the equality  $\sum_k C_{A_{k+1}}(y) \frac{w_{k+1} x^k}{w_k} = \beta(y) \sum_k C_{A_k}(y) x^k$ . But this implies that the quotient  $\frac{w_{k+1} C_{k+1}}{w_k C_k}$  equals  $\beta$  for all  $k$ , i.e., it is independent of  $k$ . As this implies  $A \in R_W$ , the proof of the theorem is now complete.  $\square$

The proofs of Proposition 2.3 and Theorem 2.4 also yield the following

**Corollary 2.5.** *Let  $\{p_n(x)\}_{n \in \mathbb{N}}$  be a Sheffer sequence, and let  $\alpha$  and  $\beta$  be the elements of  $\mathbb{F}[[y]]$  such that  $\sum_n p_n(x) \frac{y^n}{w_n}$  equals  $\alpha(y)W(x\beta(y))$ . In addition, let  $d_l \in \mathbb{F}[[h]]$  such that  $T_{h,W}(\frac{p_n}{w_n})(x)$  is  $\sum_{k=0}^n \frac{d_{n-k}(h)p_k(x)}{w_{n-k}w_k}$ . Then  $\alpha(y) = \sum_n p_n(0) \frac{y^n}{w_n}$  and  $\sum_l \frac{d_l(h)y^l}{w_l}$  equals  $W(h\beta(y))$ .*

*Proof.* Recall that  $\alpha(y)$  is just  $C_{A_0, W}(y)$ , which is defined as  $\sum_n a_{n,0} \frac{y^n}{w_n}$ . As  $a_{n,0}$  is the constant coefficient  $p_n(0)$  of  $p_n$ , this proves the first assertion. Now, We have seen that the action of  $T_{h,W}$  multiplies  $\sum_n p_n(x) \frac{y^n}{w_n} = \alpha(y)W(x\beta(y))$  by some element of  $\mathbb{F}[[h]][[y]]$ . The proof of Proposition 2.3 showed that this multiplier is  $\sum_l d_l(h) \frac{y^l}{w_l}$ , while in Theorem 2.4 we got the expression  $W(h\beta(y))$ . The second assertion follows, which completes the proof of the corollary.  $\square$

We now describe the two natural subgroups of  $R_W$  arising from its structure as a semi-direct product in terms of special Sheffer sequences.

**Definition 2.6.** *A sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  of polynomials is called an Appell sequence of weight  $W$ , or a  $W$ -Appell sequence, if its associated operator  $Q_{A,W}$  is  $D_W$ . The sequence is said to be of  $W$ -binomial type in case evaluating the image of  $\frac{p_n}{w_n}$  under  $T_{h,W}$  at  $x$  yields  $\sum_{k=0}^n \frac{p_{n-k}(h)p_k(x)}{w_{n-k}w_k}$ .*

The connection of Definition 2.6 with the previous notions is given in

**Proposition 2.7.** *A sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is  $W$ -Appell if and only if it is  $W$ -Sheffer and the polynomials  $d_{n,k,A,W}(h)$  are just  $h^{n-k}$ , i.e., its  $\beta$ -parameter is trivial.*

*Proof.* The argument proving Lemma 2.2 with  $p_n(x) = x^n$  and  $Q_{A,W} = D_W$  shows that  $D_W$  and  $T_{h,W}$  commute. Hence Appell sequences are Sheffer, and it remains to show that for a  $W$ -Sheffer sequence associated to  $A \in L$  we have  $Q_{A,W} = D_W$  if and only if the  $\beta$ -parameter is trivial. Moreover, Corollary 2.5 implies that  $W(h\beta(y))$  is  $\sum_l \frac{d_l(h)y^l}{w_l}$ , while it is clear that  $W(hy) = \sum_l \frac{h^l y^l}{w_l}$ . Hence  $\beta(y) = y$  if and only if  $d_l(h) = h^l$  for every  $l$ . Now, we recall that  $\sum_n p_n(x) \frac{y^n}{w_n} = \alpha(y) \sum_k \frac{\beta(y)^k x^k}{w_k}$  for any Sheffer sequence. Applying  $Q_{A,W}$  on the left hand side multiplies it by  $y$  (since it reduces the index of  $p_n$  and  $w_n$ , but not the power of  $y$ ), while the operation of  $D_W$  on the right hand side multiplies it by  $\beta(y)$  (by reducing the index of  $w_k$  and the power of  $x$ ). It follows that if  $Q_{A,W} = D_W$  then  $\beta(y) = y$ . Conversely, if  $\beta(y) = y$  then the same argument shows that the operation of  $Q_{A,W}$  on  $\alpha(y) \sum_k \frac{x^k y^k}{w_k}$  multiplies it by  $y$ , which is the same as sending it to  $\alpha(y) \sum_k \frac{x^{k-1} y^k}{w_{k-1}}$ . Dividing by  $\alpha(y)$  and comparing the coefficients of  $y^k$  shows that the actions on  $Q_{A,W}$  and  $D_W$  coincide on the basis  $\{\frac{x^k}{w_k}\}$  of  $\mathbb{F}[x]$ , so that  $Q_{A,W} = D_W$ . This proves the proposition.  $\square$

This connection continues with

**Proposition 2.8.** *A sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is of  $W$ -binomial type if and only if it is  $W$ -Sheffer and satisfies  $p_n(0) = \delta_{n,0}$ , i.e., if  $\alpha$ -parameter is trivial.*

The symbol  $\delta_{n,0}$  here is the Kronecker delta symbol, which equals 1 if  $n = 0$  and 0 otherwise.

*Proof.* The condition for  $W$ -binomiality means that the polynomial  $d_{n,k,A,W}$  is just  $p_{n-k}$ . As this depends only on  $n - k$ ,  $W$ -binomial sequences are  $W$ -Sheffer by Lemma 2.2. Corollary 2.5 shows that the  $\alpha$ -parameter is trivial if and only if  $p_n(0) = \delta_{n,0}$  for all  $n$ . As  $T_{0,W}$  is the identity map, substituting  $h = 0$  in Definition 2.6 and using the fact that the  $p_k$  are linearly independent shows that binomial sequences must satisfy the condition  $p_n(0) = \delta_{n,0}$ . Conversely, if  $\{p_n(x)\}_{n \in \mathbb{N}}$  is a Sheffer sequence with a trivial  $\alpha$ -parameter, then  $\sum_n p_n(x) \frac{y^n}{w_n}$  is  $W(x\beta(y))$  for some  $\beta \in \mathbb{F}[[y]]$  with  $v(\beta) = 1$ . On the other hand, Corollary 2.5 shows that the coefficient of  $\frac{y^l}{w_l}$  in  $W(h\beta(y))$  is the polynomial  $d_l(h)$  from the operation of  $T_{h,W}$ . This implies the equality  $d_n = p_n$  for every  $n$ , which completes the proof of the proposition.  $\square$

We have described several notions in terms of the behavior with respect to  $T_{h,W}$ . Many of them could have equivalently been defined using only  $D_W$ , as follows from the following



**Proposition 2.9.** *Any operator on  $\mathbb{F}[x]$  commutes with  $D_W$  if and only if it commutes with all the  $W$ -translations  $T_{h,W}$ .*

*Proof.* The coefficient of  $\frac{h^l}{w_l}$  in the image of  $\frac{x^n}{w_n}$  under  $T_{h,W}$  can be seen as the image of  $\frac{x^n}{w_n}$  under  $D_W^l$ . Hence, we may write  $T_{h,W}$ , as an operator on  $\mathbb{F}[x]$ , as  $W(hD_W) = \sum_l \frac{D_W^l}{w_l}$  (this is well-defined, since any polynomial is annihilated by a high enough power of  $D_W$ , so that the sum arising from applying the latter operator to any polynomial in  $\mathbb{F}[x]$  is essentially finite). Hence if an operator commutes with  $D_W$  then it commutes with any (finite or infinite) linear combination of its powers, and in particular with  $T_{h,W}$  for any  $h \in \mathbb{F}$ . Conversely, comparing the coefficients of  $h$  in an equation standing for the commutation of an operator with  $W(hD_W)$  (as a relation in power series in  $h$  over the space of operators on  $\mathbb{F}[x]$ ) yields the commutativity of that operator with  $D_W$ . This proves the proposition.  $\square$

Using Proposition 2.9, together with restricting attention to the coefficient of  $h$  in some power series appearing in the above arguments, proves the following assertions: The sequence of polynomials corresponding to the matrix  $A \in L$  is  $W$ -Sheffer if and only if the associated operator  $Q_{A,W}$  commutes with  $D_W$ ; This is equivalent to the coefficient  $\tilde{d}_{n,k,A,W}$  which replaces  $d_{n,k,A,W}(h)$  if  $T_{h,W}$  is replaced by  $D_W$  depending only on  $n-k$ , and to  $D_W$  multiplying  $\sum_n p_n(x) \frac{y^n}{w_n}$  by a power series in  $y$  which is independent of  $x$ . The latter multiplier is then  $\sum_l \tilde{d}_l \frac{y^l}{w_l}$ , and it equals just the  $\beta$ -parameter of the sequence.

### 3 Linear Functionals on Polynomials

We now consider the space  $\mathbb{F}[x]^*$  of linear functionals on  $\mathbb{F}[x]$ . Fix a power series  $W(t) = \sum_n \frac{t^n}{w_n} \in \mathbb{F}[[t]]$  as above.

**Lemma 3.1.** *The bilinear map taking two elements  $\varphi$  and  $\psi$  of  $\mathbb{F}[x]^*$  to the linear functional defined by  $\varphi \cdot_W \psi \left( \frac{x^n}{w_n} \right) = \sum_{k=0}^n \varphi \left( \frac{x^k}{w_k} \right) \psi \left( \frac{x^{n-k}}{w_{n-k}} \right)$  defines a product on  $\mathbb{F}[x]^*$ , making it a ring which is isomorphic to  $\mathbb{F}[[y]]$ .*

*Proof.* For any  $n \in \mathbb{N}$  we denote  $\varphi \left( \frac{x^n}{w_n} \right)$  by  $b_n$  and  $\psi \left( \frac{x^n}{w_n} \right)$  by  $c_n$ . We then identify  $\varphi$  and  $\psi$  with the elements  $\sum_n b_n y^n$  and  $\sum_n c_n y^n$  of  $\mathbb{F}[[y]]$  respectively. The fact that the elements  $\frac{x^n}{w_n}$  are a basis of  $\mathbb{F}[x]$  implies that this identification is an isomorphism of vector spaces between  $\mathbb{F}[x]^*$  and  $\mathbb{F}[[y]]$ . The fact that  $\varphi \cdot_W \psi \left( \frac{x^n}{w_n} \right) = \sum_{k=0}^n b_k c_{n-k}$  is the coefficient of  $y^n$  in the product of the latter two power series now shows that our identification preserves products as well. As all the ring axioms can now be transferred from those of  $\mathbb{F}[[y]]$ , this completes the proof of the lemma.  $\square$

In fact, the identification appearing in the proof of Lemma 3.1 can be defined by letting  $\varphi$ , as a functional on polynomials in  $x$ , operate on the expression  $W(xy) = \sum_n \frac{x^n y^n}{w_n}$ , producing an element of  $\mathbb{F}[[y]]$ . Note that this expression

is the one corresponding, as in Proposition 1.5, to the trivial element of  $R_W$ . Lemma 3.1 has the following

**Corollary 3.2.** *For  $\varphi \in \mathbb{F}[x]^*$  define  $v(\varphi) = \min \{n \in \mathbb{N} \mid \varphi(x^n) \neq 0\}$  (and  $\varphi(0) = \infty$ ). Then  $v$  is multiplicative.*

The number  $v(\varphi)$  from Corollary 3.2 is also the minimal number  $n$  such that  $\varphi(p) = 0$  wherever the degree of  $p$  is strictly smaller than  $n$ .

*Proof.* The assertion follows immediately from Lemma 3.1 and the multiplicative property of the valuation on  $\mathbb{F}[[y]]$ , since  $v(\varphi)$  coincides with the valuation of its image in  $\mathbb{F}[[y]]$ .  $\square$

Let  $\varepsilon_h \in \mathbb{F}[x]^*$  be the functional of evaluation at  $h$ , i.e.,  $\varepsilon_h(p) = p(h)$ . Then  $\varepsilon_h$  corresponds to the power series  $W(hy) \in \mathbb{F}[[y]]$ , and in particular  $\varepsilon_0$  corresponds to 1 and is the identity of  $\mathbb{F}[x]^*$ . The relation between multiplication of linear functionals and operators on polynomials is given in the following

**Lemma 3.3.** *Let  $S$  be a linear operator on  $\mathbb{F}[x]$ . Then the operation  $\varphi \mapsto \varphi \circ S$  on  $\mathbb{F}[x]^*$  is obtained through multiplication with some element  $\psi \in \mathbb{F}[x]^*$  if and only if  $S$  does not increase the degrees of polynomials and commutes with  $D_W$ . In this case  $\psi$  is uniquely determined as  $\varepsilon_0 \circ S$ . For the particular case  $S = T_{h,W}$  we get  $\psi = \varepsilon_h$ .*

*Proof.* It suffices to consider the operations on  $\frac{x^n}{w_n}$ . Write the image  $S(\frac{x^n}{w_n})$  as  $\sum_k \frac{b_{n,k,S}}{w_{n-k}} \cdot \frac{x^k}{w_k}$ . Then  $\varphi \circ S(\frac{x^n}{w_n}) = \sum_k \frac{b_{n,k,S}}{w_{n-k}} \varphi(\frac{x^k}{w_k})$  (by linearity), and we have to compare it with  $(\varphi \cdot_W \psi)(\frac{x^n}{w_n}) = \sum_{k=0}^n \psi(\frac{x^{n-k}}{w_{n-k}}) \varphi(\frac{x^k}{w_k})$  for some fixed  $\psi$ . This equality holds for every  $\varphi \in \mathbb{F}[x]^*$  if and only if  $b_{n,k,S} = \psi(\frac{x^{n-k}}{w_{n-k}})$  for every  $n$  and  $k$ . But this is equivalent to  $b_{n,k,S}$  depending only on  $n-k$  and vanishing for  $n < k$  (so that the degree of  $S(p)$  cannot exceed that of  $p$ ), together with the equality  $\psi(x^l) = b_{l,0,S}$ . The first assertion now follows as in the proof of Lemma 2.2, if we replace  $Q_{A,W}$  by  $D_W$ ,  $T_{h,W}$  by  $S$ , and  $d_{n,k,A,W}(h)$  by  $b_{n,k,S}$ . For evaluating  $\psi$  we either observe that  $\varepsilon_0 \circ S(\frac{x^n}{w_n})$  equals just  $\frac{b_{n,0,S}}{w_n}$  (like  $\psi(\frac{x^n}{w_n})$ ) for every  $n \in \mathbb{N}$ , or use the fact that  $\varepsilon_0$  is the multiplicative identity of  $\mathbb{F}[x]^*$  to establish  $\psi = \varepsilon_0 \psi = \varepsilon_0 \circ S$ . These conditions are clearly satisfied for  $T_{h,W}$ , where its defining equation implies that  $b_{n,k,T_{h,W}} = h^{n-k}$  (indeed depending only on  $n-k$ ). Its corresponding functional thus satisfies  $\psi(x^l) = h^l = \varepsilon_h(x^l)$  for every  $l \in \mathbb{N}$ . This completes the proof of the lemma.  $\square$

From this we may deduce another characterization of Sheffer sequences, as is given in

**Theorem 3.4.** *A graded sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  of polynomials is  $W$ -Sheffer if and only if there exists another graded sequence  $\{d_l(x)\}_{l \in \mathbb{N}}$  of polynomials such that the equality  $\varphi \cdot_W \psi(\frac{p_n}{w_n}) = \sum_{k=0}^n \varphi(\frac{p_k}{w_k}) \psi(\frac{d_{n-k}}{w_{n-k}})$  holds for every  $\varphi$  and  $\psi$  from  $\mathbb{F}[x]^*$ .*

*Proof.* Assume first that the sequence  $\{d_l(x)\}_{l \in \mathbb{N}}$  exists, and consider the desired equality with  $\varphi$  arbitrary and with  $\psi$  the unique element which sends  $\frac{d_1}{w_1}$  to 1 and annihilates all the other  $d_l$ s. Then the equality  $\varphi \cdot_W \psi\left(\frac{p_n}{w_n}\right) = \varphi\left(\frac{p_{n-1}}{w_{n-1}}\right)$  holds for every  $\varphi \in \mathbb{F}[x]^*$  and  $n \in \mathbb{N}$ , and the right hand side can be written as  $(\varphi \circ Q_{A,W})\left(\frac{p_n}{w_n}\right)$ . But then Lemma 3.3 implies that  $Q_{A,W}$  must commute with  $D_W$ , so that  $\{p_n(x)\}_{n \in \mathbb{N}}$  is  $W$ -Sheffer by Proposition 2.9. Conversely, by applying Proposition 1.5 we obtain the equality

$$\sum_n \varphi\left(\frac{p_n}{w_n}\right)(x)y^n = \varphi\left(\alpha(y)W(x\beta(y))\right) = \alpha(y) \sum_l \beta^l(y) \varphi\left(\frac{x^l}{w_l}\right).$$

Replacing  $\varphi$  by a product  $\varphi \cdot_W \psi$ , we decompose the image  $\varphi \cdot_W \psi\left(\frac{x^l}{w_l}\right)$  according to the definition and write the  $l$ th power of  $\beta$  as the product of its  $k$ th and  $(l-k)$ th powers. Putting the expressions back together, the terms involving  $\varphi$  give us  $\sum_n \varphi\left(\frac{p_n}{w_n}\right)(x)y^n$  again, while the ones in which  $\psi$  appear can be written (by the same argument) as  $\sum_n \psi\left(\frac{d_n}{w_n}\right)(x)y^n$ . Here  $\{d_l(x)\}_{l \in \mathbb{N}}$  is the  $W$ -binomial sequence with the same  $\beta$ -parameter as  $\{p_n(x)\}_{n \in \mathbb{N}}$  (recall that  $\alpha(y)$  multiplies only the first expression). Comparing the coefficient of  $y^n$  in the product of these series with the initial coefficient  $\varphi \cdot_W \psi\left(\frac{p_n}{w_n}\right)(x)$  yields the desired equality. This completes the proof of the theorem.  $\square$

We turn to a yet another description of Sheffer sequences, in terms of dual bases. If a sequence  $\{\varphi_r\}_{r \in \mathbb{N}}$  of functionals in  $\mathbb{F}[x]^*$  satisfies  $v(\varphi_r) \rightarrow \infty$  as  $r \rightarrow \infty$ , then  $\sum_r \varphi_r$  (or more generally  $\sum_r c_r \varphi_r$  for any sequence  $\{c_r\}_{r \in \mathbb{N}}$  of scalars) produces a well-defined element of  $\mathbb{F}[x]^*$ . Indeed, either transfer to  $\mathbb{F}[[y]]$ , or observe that evaluating the image of any polynomial in  $\mathbb{F}[x]$  under this series involves only finitely many elements. Conversely,  $\sum_r \varphi_r$  converges in this sense if and only if this condition on  $v(\varphi_r)$  holds. In this case we shall call  $\{\varphi_r\}_{r \in \mathbb{N}}$  a *pseudo-basis* of  $\mathbb{F}[x]^*$  if any  $\eta \in \mathbb{F}[x]^*$  can be presented as  $\sum_k c_r \varphi_r$  with a unique sequence of scalars  $\{c_r\}_{r \in \mathbb{N}}$ . We shall need

**Lemma 3.5.** *Let  $\{\varphi_r\}_{r \in \mathbb{N}} \subseteq \mathbb{F}[x]^*$  be a pseudo-basis of  $\mathbb{F}[x]^*$ , and define  $b_{r,l}$  to be the uniquely determined coefficient such that  $\sum_r b_{r,l} \varphi_r$  is the element of  $\mathbb{F}[x]^*$  which takes each monomial  $\frac{x^k}{w_k}$  to  $\delta_{k,l}$ . Then the coefficient  $b_{r,l}$  vanishes, for fixed  $r$ , starting from some value of  $l$ .*

*Proof.* For any  $d$  and  $r$  in  $\mathbb{N}$  we denote  $V_{r,d}$  the space of the finite sequences  $\{c_k\}_{k=0}^r$  which admit an extension to an infinite sequence  $\{c_k\}_{k=0}^\infty$  such that  $v\left(\sum_{k=0}^\infty c_k \varphi_k\right) \leq d$ . This sequence of finite-dimensional vector spaces is decreasing with  $d$ , and we define  $U_r = \bigcap_d V_{r,d}$ . Since we consider the existence of extensions, the restriction map from sequences up to  $r+1$  to sequences up to  $r$  takes  $V_{d,r+1}$  surjectively onto  $V_{r,d}$ . As decreasing sequences of finite-dimensional spaces must stabilize,  $U_{r+1}$  surjects onto  $U_r$  as well (since these are  $V_{r+1,d}$  and  $V_{r,d}$  respectively for large enough  $d$ ).

Assume now that our lemma does not hold. As the functional  $\sum_r b_{r,l} \varphi_r$  has valuation  $l$  by definition, we find that  $\{b_{k,l}\}_{k=0}^r$  is an element of  $V_{r,d}$  for every

$l \geq d$ . If  $b_{r,l} \neq 0$  for infinitely many  $l$  then  $V_{r,d} \neq \{0\}$  for all  $d$ , and the same stabilization argument from above shows that  $U_r \neq \{0\}$  as well. Starting from a non-zero element of  $U_r$ , we choose a pre-image in  $U_{r+1}$ , and continuing in this manner we get a non-zero sequence  $\{c_k\}_{k=0}^\infty$  whose finite beginning up to  $r$  lies in  $U_r$  for any  $r \in \mathbb{N}$ . We claim that  $\sum_k c_k \varphi_k = 0$ . Indeed, given any  $d \in \mathbb{N}$  we choose  $r$  such that  $v(\varphi_k) > d$  for all  $k > r$ , and then  $\{c_k\}_{k=0}^r \in U_r \subseteq V_{r,d+1}$  admits some continuation such that  $v(\sum_{k=0}^r c_k \varphi_k + \sum_{k=r+1}^\infty \tilde{c}_k \varphi_k) > d$ . But the difference between this functional and the one under consideration is based only on  $\varphi_k$  with valuation larger than  $d$ . It follows that  $v(\sum_{k=0}^\infty c_k \varphi_k) > d$  for every  $d$ , which proves our claim. But this is a contradiction to the fact that a functional (in this case the 0 functional) has a *unique* such presentation (compare with the 0 sequence). This contradiction proves the lemma.  $\square$

Let now  $\{p_n(x)\}_{n \in \mathbb{N}}$  be an arbitrary (not necessarily graded) sequence of polynomials, which we assume to be a basis for  $\mathbb{F}[x]$ . We say that a sequence of functionals  $\{\varphi_r\}_{r \in \mathbb{N}} \subseteq \mathbb{F}[x]^*$  is the *W-dual basis* of  $\{p_n(x)\}_{n \in \mathbb{N}}$  if the equality  $\varphi_r(\frac{p_n}{w_n}) = \delta_{n,r}$  holds for every  $n$  and  $r$ . We now have

**Proposition 3.6.** *Any W-dual basis is a pseudo-basis. Conversely, any pseudo-basis  $\{\varphi_r\}_{r \in \mathbb{N}}$  of  $\mathbb{F}[x]^*$  is W-dual to a unique basis  $\{p_n(x)\}_{n \in \mathbb{N}}$  of  $\mathbb{F}[x]$ . In addition, this sequence of polynomials is graded if and only if  $v(\varphi_r) = r$  for all  $r$ , and any sequence with  $v(\varphi_r) = r$  arises in this way.*

*Proof.* Let  $\{\varphi_r\}_{r \in \mathbb{N}}$  be the sequence which is W-dual to  $\{p_n(x)\}_{n \in \mathbb{N}}$ . We need to show that  $v(\varphi_r) \rightarrow \infty$  as  $r \rightarrow \infty$ . Fix  $N \in \mathbb{N}$ . As  $\{p_n(x)\}_{n \in \mathbb{N}}$  is a basis, the monomials up to  $x^N$  are linear combinations of the  $p_n$ s. Hence all of them involve only finitely many of the latter. If  $p_M$  is the maximal one appearing in any of these monomials, then it is clear from the definition (and from linearity) that  $\varphi_r(x^n)$  vanishes for every  $r > M$  and  $n \leq N$ . Hence  $v(\varphi_r) > N$  for any such  $r$ , establishing the valuation condition. Given any  $\psi \in \mathbb{F}[x]^*$ , it is now clear that  $\sum_r \psi(\frac{p_r}{w_r}) \varphi_r$  coincides with  $\psi$  on all the  $p_n$ s. Hence the latter functional coincides with  $\psi$  since the  $p_n$  are assumed to be a basis for  $\mathbb{F}[x]$ . But if we change a coefficient, the value on some  $p_n$  will change, so that this combination is unique. This proves the first assertion.

Conversely, let  $\{\varphi_r\}_{r \in \mathbb{N}}$  be a pseudo-basis, and denote  $\varphi_r(\frac{x^k}{w_k})$  by  $s_{k,r}$ . The condition on the  $v(\varphi_r)$  shows that  $s_{k,r} = 0$  for large enough  $r$  if  $k$  is fixed, so that the matrix with entries  $s_{k,r}$  is row-finite. As  $\{\varphi_r\}_{r \in \mathbb{N}}$  is a pseudo-basis, there exist for every  $l \in \mathbb{N}$  (uniquely determined) coefficients  $b_{r,l}$  such that  $\sum_r b_{r,l} \varphi_r$  is the element of  $\mathbb{F}[x]^*$  which takes each monomial  $\frac{x^k}{w_k}$  to  $\delta_{k,l}$ . This translates to the equality  $\sum_r s_{k,r} b_{r,l} = \delta_{k,l}$ . As Lemma 3.5 shows that the matrix  $B$  with entries  $b_{r,l}$  is also row-finite, both these matrices may be seen as representing linear operators on a vector space with a countable basis, whose product  $SB$  is the identity. But as  $S$  is invertible (it represents a bijective operator), we find that  $B = S^{-1}$  hence  $BS$  is the identity as well. For each  $n \in \mathbb{N}$  we now define  $p_n(x) = w_n \sum_k b_{n,k} \frac{x^k}{w_k}$ , which is a polynomial by Lemma 3.5. It follows that  $\varphi_r(\frac{p_n}{w_n})$  equals  $\sum_k b_{n,k} s_{k,r}$ , which was seen to be just  $\delta_{n,r}$ . The  $p_n$  are therefore

linearly independent, and the inverse relations between  $S$  and  $B$  shows that the equality  $\frac{x_l}{w_l} = \sum_n s_{l,n} \frac{p_n}{w_n}$  holds for every  $l \in \mathbb{N}$ . It follows that the  $p_n$ s form a basis for  $\mathbb{F}[x]$ , and  $\{\varphi_r\}_r$  is the  $W$ -dual basis. This proves the second assertion.

For the third assertion, note that if  $\{p_n(x)\}_{n \in \mathbb{N}}$  is graded then  $\{p_n(x)\}_{n=0}^l$  span the space of polynomials of degree at most  $l$ . Hence  $\varphi_r$  vanishes on each polynomial of degree smaller than  $r$ , but not on  $p_r$  of degree  $r$ , showing that  $v(\varphi_r) = r$ . Conversely, if  $v(\varphi_r) = r$  then the matrix  $S$  from the previous paragraph lies in the group  $L$  from Proposition 1.2. Hence so does its inverse  $C$ , showing that the corresponding sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is graded (and in particular forms a basis for  $\mathbb{F}[x]$ ). This completes the proof of the theorem.  $\square$

Proposition 3.6 shows that sequences  $\{\varphi_r\}_{r \in \mathbb{N}}$  with  $v(\varphi_r) = r$  are precisely those which are dual to graded sequences. Corollary 3.2 shows that wherever  $v(\xi) = 0$  and  $v(\eta) = 1$  the sequence in which  $\varphi_r = \xi \cdot_W \eta_W^r$  (where  $\eta_W^r$  stands for the  $r$ th power of  $\eta$  in the product associated to  $W$  as in Lemma 3.1) has the latter property. We now prove

**Theorem 3.7.** *The graded sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is  $W$ -Sheffer if and only if its  $W$ -dual basis takes the form  $\varphi_r = \xi \cdot_W \eta_W^r$  for some elements  $\xi$  and  $\eta$  of  $\mathbb{F}[x]^*$  with  $v(\xi) = 0$  and  $v(\eta) = 1$ .*

*Proof.* By Theorem 2.4,  $\{p_n(x)\}_{n \in \mathbb{N}}$  is Sheffer if and only if the corresponding matrix  $A$  lies in  $R_W$ . By Proposition 1.4 this is equivalent to its inverse lying in  $R_W$ . Now, consider the matrix  $S$  in which  $s_{k,r} = \varphi_r(\frac{x^k}{w_k})$ . On the one hand, the proof of Proposition 3.6 shows that this is the matrix which is inverse to  $B$  in which  $b_{n,k} = \frac{a_{n,k} w_k}{w_n}$ . Hence the  $(k,r)$ -entry of  $A^{-1}$  is  $\frac{w_k s_{k,r}}{w_r}$ . It follows that the sum  $\sum_k s_{k,r} y^k$  forms, on the one hand, the series  $C_{A_r^{-1}, W}(y)$  associated to the element  $A^{-1} \in L$ , multiplied by  $w_r$ . On the other hand, it is the image of  $\sum_k \frac{x^k y^k}{w_k} = W(xy)$  under  $\varphi_r$ , i.e., the element of  $\mathbb{F}[[y]]$  which is associated to  $\varphi_r$ . Since this association preserves products of power series, it follows that  $\varphi_r$  is of the form  $\xi \cdot_W \eta_W^r$  for some  $\xi$  and  $\eta$  as above if and only if the columns of  $A^{-1}$  satisfy the definition of columns of elements of  $R_W$ . This proves the theorem.  $\square$

Theorems 3.4 and 3.7 are related, as follows. If  $\{p_n(x)\}_{n \in \mathbb{N}}$  is  $W$ -Sheffer, then consider the sequence of polynomials  $\{d_l(x)\}_{l \in \mathbb{N}}$  from Theorem 3.4. We take  $\eta$  to be the functional sending  $\frac{d_1}{w_1}$  to 1 and the other  $d_l$ s to 0 (this functional appears in the proof of one direction there), and let  $\xi$  be the one taking  $p_0$  to 1 and the other  $p_n$ s to 0. Extending the proof of Theorem 3.4 by a simple induction then shows that  $\xi \cdot_W \eta_W^r$  sends  $\frac{p_n}{w_n}$  to 0 if  $n < r = v(\xi \cdot_W \eta_W^r)$  and to  $\xi(\frac{p_{n-r}}{w_{n-r}}) = \delta_{n,r}$  otherwise. This is one description for the basis which is  $W$ -dual to a  $W$ -Sheffer sequence. On the other hand,  $\xi$  and  $\eta$  are characterized by taking  $W(xy)$  to the  $\alpha$  and  $\beta$  parameters of  $A^{-1}$  respectively, as the proof of Theorem 3.4 shows. We also remark that several of the results of this section may also be established using the fact that the  $W$ -dual basis element  $\varphi_r$  is characterized by the equality  $\varphi_r(\sum_n p_n(x) \frac{y^n}{w_n}) = y^r$  (which easily follows from the definition), which for a  $W$ -Sheffer sequence becomes  $\varphi_r[\alpha(y)W(x\beta(y))] = y^r$  (by Proposition 1.5).

## 4 Group Operations and Conjugate Subgroups

We can consider the operators  $D_W$ ,  $T_{h,W}$ , and  $Q_{A,W}$  as infinite lower triangular matrices as well. Indeed, the formula defining the action of  $D_W$  shows that it coincides with the operation of the matrix, which we denote  $M_W$ , whose  $(n, k)$ -entry is  $\frac{w_n}{w_{n-1}}$  if  $k = n - 1$  and 0 otherwise. This allows us to give another characterization of the  $W$ -Appell sequences, as in

**Proposition 4.1.** *The sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$ , corresponding to  $A \in L$ , is  $W$ -Appell if and only if the matrix  $A$  can be obtained by substituting  $y = M_W$  inside a power series from  $\mathbb{F}[[y]]$  with valuation 0. Explicitly, we have  $A = \alpha(M_W)$  with the  $\alpha$ -parameter of  $A \in R_W$ .*

Note that as every power of  $M_W$  is supported on a different diagonal, we can substitute  $M_W$  inside power series, and there is no problem of convergence.

*Proof.* Evaluating the powers of  $M_W$  shows that  $(M_W^r)_{n,k} = \frac{w_n}{w_{n-r}} \delta_{k,n-r}$  (using the Kronecker  $\delta$  symbol again) for any  $r \in \mathbb{N}$ . Hence multiplying each power by some element of  $\mathbb{F}$ , which we write as  $\frac{b_r}{w_r}$ , and summing over  $r$ , yields a matrix whose  $(n, k)$ -entry is  $\frac{b_{n-k} w_n}{w_k w_{n-k}}$  (and 0 if  $n < k$ ). Under the valuation assumption, namely  $b_0 \neq 0$ , this matrix is also invertible, hence lies in  $L$ . On the other hand, Proposition 2.7 shows that  $\{p_n(x)\}_{n \in \mathbb{N}}$  is  $W$ -Appell if and only if the expression  $\sum_n p_n(x) \frac{y^n}{w_n}$  equals  $\alpha(y)W(xy)$  for some  $\alpha \in \mathbb{F}[[y]]$  with  $v(\alpha) = 0$ . Expanding both sides yields

$$\sum_n \sum_{k=0}^n a_{n,k} \frac{x^k y^n}{w_n} = \sum_n p_n(x) \frac{y^n}{w_n} = \alpha(y)W(xy) = \sum_l \frac{a_{l,0} y^l}{w_l} \sum_k \frac{x^k y^k}{w_k}$$

(since  $\alpha = C_{A_0,W}$ ). From this we deduce, by comparing the coefficients of  $x^k y^n$ , that a necessary and sufficient condition for  $\{p_n(x)\}_{n \in \mathbb{N}}$  to be  $W$ -Appell is that the equality  $a_{n,k} = \frac{a_{n-k,0} w_n}{w_k w_{n-k}}$  holds for every  $n$  and  $k$ . But the right hand side of the latter equality was seen to be the  $(n, k)$ -entry of the matrix  $\sum_l \frac{a_{l,0}}{w_l} M_W^l$ . This proves the first assertion, while the second one follows directly from the fact that  $\alpha(y)$  is  $C_{A_0,W}(y) = \sum_l a_{l,0} \frac{y^l}{w_l}$ . This proves the proposition.  $\square$

From Proposition 4.1 we directly deduce

**Corollary 4.2.** *The operator  $T_{h,W}$  corresponds to the matrix in  $R_W$  with  $\alpha$ -parameter  $\alpha(y) = W(hy)$  and trivial  $\beta$ -parameter.*

*Proof.* The proof of Proposition 2.9 shows that  $T_{h,W} = W(hD_W)$  as operators on  $\mathbb{F}[x]$ . Hence it is represented by the matrix which is the result of substituting  $y = M_W$  in the series  $W(hy)$ . The corollary now follows from the second assertion of Proposition 4.1.  $\square$

We now show that many the previous notions can be described in terms of a certain group action of  $L$ . We have seen that the operators  $D_W$  can be represented by matrices. The same can be done with  $Q_{A,W}$ . In fact, the main idea for this presentation is given in the following

**Lemma 4.3.** *The operator  $Q_{A,W}$  is described by the matrix  $A^{-1}M_WA$ .*

*Proof.* Let  $\tilde{Q}$  be the matrix representing the operation of  $Q_{A,W}$  on the powers of  $x$ . Then the sequence  $\{Q_{A,W}(x^n)\}_{n \in \mathbb{N}}$  is  $\tilde{Q}$  times the monomial sequence. The linearity of  $Q_{A,W}$  shows that sequence  $\{Q_{A,W}(p_n)(x)\}_{n \in \mathbb{N}}$  can be written as  $A\tilde{Q}$  times the monomial sequence, and multiplying from the right by the row vector from the proof of Proposition 1.4 yields the image of  $\sum_n p_n(x) \frac{y^n}{w_n}$  under  $Q_{A,W}$ . But we have seen in the proof of Proposition 2.7 that this action multiplies it by  $y$ . On the other hand, multiplying that row vector by  $M_W$  from the right also have the same effect of multiplying by the scalar  $y$ . As  $\{p_n\}_{n \in \mathbb{N}}$  is  $A$  times the monomial sequence, we find that  $y \sum_n p_n(x) \frac{y^n}{w_n}$  can also be presented as the row vector from the proof of Proposition 1.4 times  $M_WA$  times the monomial sequence. The fact that the products of different powers of  $x$  and  $y$  are linearly independent implies that if putting two matrices between the row vector from the proof of Proposition 1.4 and the monomial sequence yields the same result then the matrices must be equal. As this is the case for  $A\tilde{Q}$  and  $M_WA$ , the proof of the lemma is now complete.  $\square$

To make the proof of Lemma 4.3 a bit more visible, we note that the explicit meaning of the equality  $Q_{A,W}(p_n) = \frac{w_n p_{n-1}}{w_{n-1}}$  is

$$\sum_{l=0}^n a_{n,l} Q_{A,W}(x^l) = Q_{A,W}(p_n)(x) = \frac{w_n p_{n-1}}{w_{n-1}} = \sum_{k=0}^{n-1} \frac{w_n a_{n-1,k} x^k}{w_{n-1}}.$$

We write the coefficient  $\frac{w_n a_{n-1,k}}{w_{n-1}}$  of  $x_k$  as  $\sum_r (M_W)_{n,r} a_{r,k}$ , and observe that the coefficient of  $x^k$  is now the  $(n,k)$ -entry of  $M_WA$  on the right hand side and the  $(n,k)$ -entry of  $A\tilde{Q}$  on the left hand side. Hence the desired equality indeed follows.

In view of Lemma 4.3 we consider the right action of  $L$  on the set of strictly lower triangular matrices (i.e., those lower triangular matrices in which all the entries on the main diagonal also vanish), in which  $A \in L$  takes a strictly lower triangular matrix  $M$  to  $A^{-1}MA$ . Using this operation we obtain an alternative description of the groups of  $W$ -Appell and  $W$ -Sheffer sequences, as in

**Theorem 4.4.** *The group of  $W$ -Appell sequences is the stabilizer of  $M_W$  in  $L$ . The group  $R_W$  of  $W$ -Riordan arrays (or  $W$ -Sheffer sequences) is the normalizer of the latter group in  $L$ .*

*Proof.* The first assertion follows directly from Definition 2.6 and Lemma 4.3. For the second assertion we first observe that Corollary 4.2 implies that the operation of  $T_{h,W}$  on the monomial sequence is via the matrix which we may write as  $W(hM_W)$ . The commutation of the operators  $T_{h,W}$  and  $Q_{A,W}$  translates, via the latter assertion and Lemma 4.3, to the commutation of  $A^{-1}M_WA$  and  $W(hM_W)$ . This is equivalent to the assertion that  $AW(hM_W)A^{-1}$  and  $M_W$  commute. Comparing the powers of  $h$ , we find that the latter condition holds if and only if  $M_W$  commutes with  $AM_W^l A^{-1}$  for every  $l \in \mathbb{N}$  (alternatively,  $Q_{A,W}$

commutes with  $D_W$  by Proposition 2.9, so that  $M_W$  commutes with  $AM_WA^{-1}$ , hence with all its powers). Multiplying each such term by some coefficient, where the coefficient of the 0th power is non-zero, and taking the sum thus obtained, shows that the condition in question is equivalent to the assertion that  $ABA^{-1}$  commutes with  $M_W$  for every  $B \in L$  which corresponds to a  $W$ -Appell sequence (by Proposition 4.1). But Lemma 4.3 and the first assertion here translate the latter condition to the statement that conjugation by  $A$  takes  $W$ -Appell sequences to  $W$ -Appell sequences (if one identifies sequences with elements of  $L$  as in Proposition 1.1). The proof of the theorem is now complete.  $\square$

Theorem 4.4 gives us a way to relate the Riordan groups of different weights, as in

**Corollary 4.5.** *Let  $W(t) = \sum_l \frac{t^n}{w_n}$  and  $\widetilde{W}(t) = \sum_l \frac{t^n}{\widetilde{w}_n}$  be two power series in  $\mathbb{F}[[t]]$ , with  $w_0 = \widetilde{w}_0 = 1$  and  $w_n \widetilde{w}_n \neq 0$  for each  $n \in \mathbb{N}$ . Then the groups  $R_W$  and  $R_{\widetilde{W}}$  are conjugate in  $L$ , and this conjugation takes the subgroup of  $W$ -Appell sequences to the subgroup of  $\widetilde{W}$ -Appell sequences.*

*Proof.* Let  $U$  be the diagonal element of  $L$  whose  $n$ th diagonal entry is  $\frac{w_n}{\widetilde{w}_n}$ . Then a simple calculation shows that  $U^{-1}M_WU = M_{\widetilde{W}}$ , i.e., the action of  $U$  from Lemma 4.3 takes  $M_W$  to  $M_{\widetilde{W}}$ . But this implies that conjugation by  $U$  takes the stabilizer of  $M_W$  to the stabilizer of  $M_{\widetilde{W}}$ . It now follows that conjugation by  $U$  takes the normalizer of the former stabilizer to normalizer of the latter stabilizer. The corollary hence follows from Theorem 4.4.  $\square$

In fact, one may obtain a more precise assertion than Corollary 4.5, as one sees in the following

**Proposition 4.6.** *The conjugation by the matrix  $U$  from the proof of Corollary 4.5 takes the element  $A \in R_W$  with parameters  $\alpha$  and  $\beta$  to the element of  $R_{\widetilde{W}}$  having the same parameters.*

Recall that as the action from Lemma 4.3 is from the right, the conjugation in Proposition 4.6 takes  $A$  to  $U^{-1}AU$ .

*Proof.* We have seen in the proof of Proposition 1.4 that multiplying  $A$  by the row vector from that proposition (based on  $W$ ) yields the row vector whose  $k$ th entry is  $\alpha(y) \frac{\beta(y)^k}{w_k}$ . Now, one easily sees that multiplying the row vector based on  $\widetilde{W}$  by  $U^{-1}$  yields the one which is based on  $W$ . Hence multiplying the former vector by  $U^{-1}A$  produces the vector with the  $\alpha(y) \frac{\beta(y)^k}{w_k}$  entries. But multiplying the latter vector by  $U$  turns the entries to  $\alpha(y) \frac{\beta(y)^k}{\widetilde{w}_k}$ . Applying the proof of Proposition 1.4 once more now proves the desired assertion.  $\square$

It follows from Proposition 4.6 that the subgroup of  $W$ -binomial sequences is also preserved under this conjugation (see Proposition 2.8, as well as Proposition 2.8 for re-proving this for Appell sequences).



The matrices  $M_W$ , for varying  $W$ , contain only one diagonal of non-zero entries. However, the operation of  $L$  produces much more matrices. Hence we define a *degree decreasing operator* to be any linear operator  $Q$  on  $\mathbb{F}[x]$  such that the degree of  $Q(p)$  is one less than the degree of  $p$  (so that in particular,  $Q$  annihilates scalars). The same argument proving Proposition 1.1 shows that these operators are in one-to-one correspondence with strictly lower triangular matrices  $M$  none of whose entries with indices  $(n, n-1)$  for  $n \in \mathbb{N}$  vanish. It is clear that the action of  $L$  preserves this set. In addition, Corollary 4.5 and Proposition 4.6 extend to

**Proposition 4.7.** *The action of  $L$  on the set of degree decreasing operators is transitive. Moreover, the set of elements  $A \in L$  which satisfy  $a_{n,0} = \delta_{n,0}$  is a subgroup of  $L$ , with respect to which the set of degree decreasing operators is a principal homogenous space.*

We recall that a *principal homogenous space* for a group  $G$  is a set on which  $G$  operates transitively with trivial stabilizers.

*Proof.* The fact that this subset of  $L$  is a subgroup is easily verified by matrix multiplication. We fix one degree decreasing operator, the one corresponding to the matrix  $M_W$  for  $W(t) = \frac{1}{1-t} = \sum_n t^n$  (i.e., with  $w_n = 1$  for all  $n$ ). It suffices to show that for any strictly lower triangular matrix  $M$ , with entries representing a degree decreasing operator there exists a unique element  $A \in L$  such that  $AM_W A^{-1} = M$  and  $a_{n,0} = \delta_{n,0}$ . We thus compare the entries of  $AM_W$  and  $MA$ . The  $(n, k)$ -entry of the former product is just  $a_{n,k+1}$ , while the same entry of the latter one is  $\sum_{l=k}^{n-1} m_{n,l} a_{l,k}$ . Comparing the two expressions shows that the equality  $AM_W = MA$  determines the columns of  $A$  by induction from the basis  $k = 0$ . Hence such  $A$  always exists, and simple induction using the same comparison shows that  $a_{n,k} = 0$  wherever  $n < k$ , and  $a_{n,n} \neq 0$  if  $a_{0,0} \neq 0$ . Thus any choice of  $\{a_{n,0}\}_{n=0}^{\infty}$  with  $a_{0,0}$  generates a unique matrix  $A \in L$  satisfying this equality, and in particular this is the case if  $a_{n,0} = \delta_{n,0}$  for any  $n$ . This completes the proof of the proposition.  $\square$

Recall that the product on  $\mathbb{F}[x]^*$  appearing in Lemma 3.1 is also based on  $W$ . We believe that there should be an action of  $L$  on the set of multiplications on  $\mathbb{F}[x]^*$  making it isomorphic to  $\mathbb{F}[[y]]$  and preserving valuations, such that the stabilizer of the multiplication  $\cdot_W$  is again the subgroup of  $L$  corresponding to  $W$ -Appell sequences. However, as such multiplications are, in some sense, 3-dimensional objects (i.e., are represented by algebraic objects whose entries have 3 indices), we do not pursue this subject here further. In addition, recalling that a polynomial sequence if  $W$ -Sheffer implies many combinatorial properties of this sequence, we conclude this section by suggesting that higher Sheffer sequences, defined by taking the normalizer in  $L$  of  $R_W$  and repeating this construction, may also turn out to have some combinatorial importance as well. We leave this question, however, for further research.

## 5 Some Examples and Relations

We present the most classical and natural examples for the choice of  $W$ , with the resulting operators.

**Example: Exponentials.** Assume the  $\mathbb{F}$  is of characteristic 0, fix  $0 \neq \lambda \in \mathbb{F}$ , and take  $w_n = \lambda^n n!$ . In this case we have  $W(t) = e^{t/\lambda}$ , and the operator  $D_W$ , taking  $\frac{x^n}{\lambda^n n!}$  to  $\frac{x^{n-1}}{\lambda^{n-1}(n-1)!}$ , is the usual derivative  $\frac{d}{dx}$  multiplied by  $\lambda$ . The  $W$ -translation  $T_{h,W}$  takes  $x^n$  to  $\sum_{k=0}^n \binom{n}{k} h^{n-k} x^k = (x+h)^n$  (in correspondence with the description of  $T_{h,W}$  as  $e^{hD_W/\lambda} = e^{hd/dx}$ ). It is thus indeed a translation  $T_{h,W}(p)(x) = p(x+h)$  (whence the name). Hence the corresponding Appell sequences are those sequences  $\{p_n(x)\}_{n \in \mathbb{N}}$  which satisfy  $p'_n = np_{n-1}$  (the  $\lambda$ s cancel). The Sheffer sequences are defined by the operator  $Q_A$  which sends  $p_n$  to  $\lambda np_{n-1}$  commuting with replacing the argument  $x$  by  $x+h$ , and the binomial sequences satisfy  $p_n(x+h) = \sum_{k=0}^n \binom{n}{k} p_{n-k}(h)p_k(x)$  (indeed, a binomial relation). This is the only case where the composition  $T_{g,W} \circ T_{h,W}$  gives just  $T_{g+h,W}$  and the product of two evaluation functionals is also an evaluation functional: Indeed, here we have  $\varepsilon_h \cdot_W \varepsilon_g = \varepsilon_{g+h}$  for any  $g$  and  $h$  from  $\mathbb{F}$ .

**Example: Power Reduction.** Now let  $\mathbb{F}$  be arbitrary, fix again such  $\lambda$ , and consider the case where  $w_n = \lambda^n$ . The series  $W(t)$  here equals  $\frac{\lambda}{\lambda-t}$ , and the operator  $D_W$ , whose action sends  $\frac{x^n}{\lambda^n}$  to  $\frac{x^{n-1}}{\lambda^{n-1}}$ , is given by  $p \mapsto \lambda \frac{p(x)-p(0)}{x}$ . The formula for  $T_{h,W}(x^n)$  is here  $\sum_{k=0}^n h^{n-k} x^k = \frac{x^{n+1}-h^{n+1}}{x-h}$ . As  $D_W$  is some normalization of  $\frac{\lambda}{x}$ , we find see that  $T_{h,W}$  roughly multiplies each polynomial by  $\frac{1}{1-h/x} = \frac{x}{x-h}$ : Indeed, the exact formula for the action of  $T_{h,W}$  here is given by  $T_{h,W}(p)(x) = \frac{xp(x)-hp(h)}{x-h}$ , which equals  $p(x) + h\tilde{p}_h(x)$  where  $\tilde{p}_h$  is a polynomial of degree one less than  $p$ . A sequence  $\{p_n(x)\}_{n \in \mathbb{N}}$  is Appell in this context if it satisfies the condition  $p_n(x) - p_n(0) = xp_{n-1}(x)$ . For Sheffer sequences we require the condition that for every  $h \in \mathbb{F}$  the operator  $Q_A$  taking  $p_n$  to  $\lambda p_{n-1}$  also takes  $\tilde{p}_{n,h}$  to  $\lambda \tilde{p}_{n-1,h}$ . The binomial sequences in this setting are those which satisfy  $p_0(x) = 1$ ,  $p_n(0) = 0$  for all  $n \geq 1$ , and the equality  $\frac{p_n(x)}{x} - \frac{p_n(h)}{h} = (x-h) \sum_{k=1}^{n-1} \frac{p_{n-k}(h)}{h} \cdot \frac{p_k(x)}{x}$  for every  $n \geq 1$ .

The choice of  $w_n = \left(\frac{\lambda}{1-q}\right)^n \prod_{j=1}^n (1-q^j)$ , with  $q \in \mathbb{F}$  which is neither 0 nor a root of unity (e.g., some number which is transcendental over the prime field of  $\mathbb{F}$ ) gives the  $q$ -umbral calculus. Here  $D_W(p)(x) = \lambda \frac{p(qx)-p(x)}{qx-x}$  is the  $q$ -derivative multiplied by  $\lambda$ , the equality which characterizes  $W$ -Appell sequences is  $p_n(qx) = p_n(x) + x(q^n-1)p_{n-1}(x)$ , and  $T_{h,W}(p)(x)$  is described by the formula  $\sum_l \sum_{k=0}^l \binom{l}{k}_q \frac{(-1)^{l-k} h^l p(q^k x)}{(qx-x)^l} q^{-k(l-k)-\binom{k}{2}}$ , where  $\binom{l}{k}_q$  is the  $q$ -binomial coefficient  $\prod_{j=k+1}^l (1-q^j) / \prod_{j=1}^{l-k} (1-q^j)$ . In this case the sequence of polynomials whose associated matrix represents  $T_{h,W}$  take the form  $p_n(x) = \prod_{j=0}^{n-1} (x + hq^j)$ , i.e., the roots of  $p_n$  are the first  $n$  terms of a geometric sequence with quotient  $q$ . As neither the formula for  $T_{h,W}$  nor the other related ones seem to reduce to succinct expressions, we do not follow the detail of this example here. We just remark that when  $\mathbb{F}$  is of characteristic 0 then the limit  $q \rightarrow 1$  exists (at least

formally), yielding the exponential example from above.

The behavior of these examples with respect to the parameter  $\lambda$  illustrate the operation of multiplying  $w_n$  by  $\lambda^n$  in the general case. This is given in the following

**Proposition 5.1.** *Let  $W(t) = \sum_n \frac{t^n}{w_n} \in \mathbb{F}[[t]]$  and a non-zero element  $\lambda \in \mathbb{F}$  be given. Then multiplying each  $w_n$  by  $\lambda^n$  leaves  $R_W$ , as well as its subgroups of Appell and binomial sequence, invariant.*

*Proof.* This operation multiplies  $D_W$  by  $\lambda$  and replaces  $W(t)$  by  $W(\frac{t}{\lambda})$ . Hence  $T_{h,W}$  remains the same. This immediately proves the assertion for the Appell sequences, as well as for Sheffer sequences since multiplying  $Q_{A,W}$  by a scalar does not affect its commutation with  $T_{h,W}$ . For sequences of binomial type this follows directly from the fact that this operation leaves the quotient  $\frac{w_n}{w_{n-k}w_k}$  invariant for every  $n$  and  $k$ . This proves the proposition.  $\square$

Note that the parameters  $\alpha$  and  $\beta$  of a Sheffer sequence are not preserved by the operation from Proposition 5.1. Indeed, the expression yielding these parameters is now  $\sum_n p_n(x) \frac{y^n}{\lambda^n w_n}$  rather than  $\sum_n p_n(x) \frac{y^n}{w_n}$ , so that this expression is  $\alpha(\frac{y}{\lambda}) W(x\beta(\frac{y}{\lambda}))$ . Moreover, we have  $W(t) = \widetilde{W}(\lambda t)$  with the relevant power series  $\widetilde{W}(t) = \sum_n \frac{t^n}{\lambda^n w_n}$  from  $\mathbb{F}[[t]]$ . Indeed, Proposition 4.6 shows that in order to preserve these parameters we need to conjugate by the diagonal matrix containing the powers of  $\lambda$  on the diagonal. Since this operation replaces the polynomial  $p_n(x)$  by  $\lambda^n p_n(\frac{x}{\lambda})$ , this is indeed the operation which yields the power series with the same  $\alpha$  and  $\beta$  for the weight  $\widetilde{W}$ .

We remark that Theorem 4.4 allows us to extend the definitions of  $R_W$ , as well as its Appell and binomial subgroups, to the case where the matrix  $M_W$  is replaced by any strictly lower diagonal matrix  $M$  with non-vanishing entries just below the main diagonal. Indeed, the Appell subgroup is the stabilizer of  $M$ , the group of corresponding Riordan arrays is its normalizer, and the binomial subgroup consists of those elements of the latter group in which  $a_{n,0} = \delta_{n,0}$ . One example for an interesting operator corresponding to such a matrix is, in characteristic 0, the *finite difference operator*  $\Delta_a$  defined by  $\Delta_a(p)(x) = p(x+a) - p(x)$  (in fact, weighted finite difference operators of the sort taking  $p$  to  $T_{a,W}(p) - p$ , in any characteristic also produce such matrices). Moreover, Proposition 4.7 shows that all these groups are again conjugate (hence isomorphic), where for preserving the group of sequences of binomial type we restrict attention to conjugators also satisfying the condition  $a_{n,0} = \delta_{n,0}$ . Even though such more general Riordan arrays represent sequences of polynomials satisfying more complicated relations, it is interesting to know that these groups are all algebraically isomorphic. The question whether every element of  $L$  lies in a subgroup of such more general Riordan arrays is also worth investigating.

The fact that a sequence of polynomials is a Sheffer sequence for some  $W(t) = \sum_n \frac{t^n}{w_n} \in \mathbb{F}[[t]]$  yields a lot of combinatorial information about the sequence. Hence it may be worthwhile to find when such a sequence is Sheffer

for two different such weights. For this we recall the *extended binomial coefficients*, defined over any field  $\mathbb{F}$  of characteristic 0 by noting that the expression  $\frac{1}{n!} \prod_{j=0}^{n-1} (\xi - j)$  for the binomial coefficient  $\binom{\xi}{n}$  makes sense for  $\xi \in \mathbb{F}$ . We shall make use of the formula given in the following

**Lemma 5.2.** *Given  $\xi$  and  $\eta$  in  $\mathbb{F}$  and  $n \in \mathbb{N}$ , we have  $\sum_{r+s=n} \binom{\xi}{r} \binom{\eta}{s} = \binom{\xi+\eta}{n}$ .*

*Proof.* We apply induction on  $n$ . The case  $n = 0$  is trivial. Assume that the equality holds for  $n$ , and write  $\binom{\xi+\eta}{n+1}$  as  $\frac{\xi+\eta-n}{n+1} \binom{\xi+\eta}{n}$ . The induction hypothesis allows us to write the latter expression as

$$\begin{aligned} \frac{\xi+\eta-n}{n+1} \sum_{r+s=n} \binom{\xi}{r} \binom{\eta}{s} &= \sum_{r+s=n} \left[ \frac{\xi-r}{n+1} + \frac{\eta-s}{n+1} \right] \binom{\xi}{r} \binom{\eta}{s} = \\ &= \sum_{s < \tilde{r}+s=n+1} \frac{\tilde{r}}{n+1} \binom{\xi}{\tilde{r}} \binom{\eta}{s} + \sum_{r < \tilde{s}+s=n+1} \frac{\tilde{s}}{n+1} \binom{\xi}{r} \binom{\eta}{\tilde{s}}, \end{aligned}$$

where we have set  $\tilde{r} = r + 1$  and  $\tilde{s} = s + 1$  respectively. The multiplying coefficients allow us to omit the strict inequalities in the summation, and by renaming the summation indices again and noting the condition on their sum we get the desired expression  $\sum_{r+s=n+1} \binom{\xi}{r} \binom{\eta}{s}$ . This proves the lemma.  $\square$

The usual binomial rule follows as a

**Corollary 5.3.** *The equality  $\binom{\xi+1}{l} = \binom{\xi}{l} + \binom{\xi}{l-1}$  holds for any  $l \geq 1$  and  $\xi \in \mathbb{F}$ .*

*Proof.* The assertion follows from the case  $\eta = 1$  in Lemma 5.2, since  $\binom{1}{s}$  vanishes for every  $s > 1$  and  $\binom{1}{0} = \binom{1}{1} = 1$ .  $\square$

Now, if one sequence is Appell then we have the following

**Theorem 5.4.** *Take  $W$  as above, and let  $A$  be the element of  $R_W$  representing a  $W$ -Appell sequence with parameter  $\alpha(y) = \sum_n c_n y^n$ , with  $c_0 \neq 0$ . This matrix lies in  $R_{\widetilde{W}}$  for some weight element  $\widetilde{W}(t) = \sum_n \frac{t^n}{\widetilde{w}_n} \in \mathbb{F}[[t]]$  in the following two cases: (i) The expression  $\gamma_k = \frac{\widetilde{w}_k w_{k+1}}{\widetilde{w}_{k+1} w_k}$  is a non-zero constant, independent of  $k$ . (ii)  $\mathbb{F}$  has characteristic 0,  $\gamma_k$  is a non-constant, never vanishing linear function of  $k$ , and  $\alpha(y) = c_0 e^{hy}$  for some  $h \in \mathbb{F}$ . Conversely, if  $c_1 \neq 0$  then  $A \in R_{\widetilde{W}}$  only if one of the conditions (i) or (ii) is satisfied.*

*Proof.* The proof of Proposition 2.7 shows that the entries of the matrix  $A$  satisfy the equality  $\sum_{n,k} a_{n,k} \frac{x^k y^n}{w_n} = \sum_l c_l \sum_k \frac{x^k y^{k+l}}{w_k}$ . It follows that  $a_{n,k} = \frac{w_n c_{n-k}}{w_k}$ . Given  $\widetilde{W}$ , we thus have  $C_{A_k, \widetilde{W}}(y) = \sum_n \frac{w_n c_{n-k} y^n}{w_k \widetilde{w}_n}$ . By Definition 1.3, we have to check when does the equality  $\widetilde{w}_k^2 C_{A_k, \widetilde{W}}^2 = \widetilde{w}_{k-1} C_{A_{k-1}, \widetilde{W}} \widetilde{w}_{k+1} C_{A_{k+1}, \widetilde{W}}$  hold for every  $k \geq 1$ . Explicitly, this equality becomes

$$\frac{\widetilde{w}_k^2}{w_k^2} \sum_{n,m} \frac{w_n w_m c_{n-k} c_{m-k} y^{n+m}}{\widetilde{w}_m \widetilde{w}_n} = \frac{\widetilde{w}_{k+1} \widetilde{w}_{k-1}}{w_{k-1} w_{k+1}} \sum_{n,m} \frac{w_n w_m c_{n-1-k} c_{m+1-k} y^{n+m}}{\widetilde{w}_m \widetilde{w}_n}.$$

Given a number  $p \geq 2k$ , the coefficients of  $y^p$  are

$$\sum_{n+m=p} c_{n-k} c_{m-k} \prod_{j=k}^{n-1} \gamma_j \prod_{j=k}^{m-1} \gamma_j \quad \text{and} \quad \sum_{n+m=p} c_{n-1-k} c_{m+1-k} \prod_{j=k+1}^{n-1} \gamma_j \prod_{j=k-1}^{m-1} \gamma_j,$$

where coefficients with negative indices are understood to be 0 while empty products equal 1. Now, if  $\gamma_j$  is some non-zero constant  $\lambda$  then this equality clearly holds, regardless of  $\alpha$  (indeed, one easily verifies that this condition on  $\gamma_j$  is equivalent to  $\widetilde{W}$  being related to  $W$  in the manner described in Proposition 5.1, making the assertion here a consequence of that proposition). Hence  $A \in R_{\widetilde{W}}$  in case (i). On the other hand, if  $\mathbb{F}$  has characteristic 0 and  $\gamma_j = \lambda - \sigma j$  for some  $\lambda$  and  $\sigma$  from  $\mathbb{F}$  with  $\sigma \neq 0$  then any product of the form  $\prod_{j=\kappa}^{\nu-1} \gamma_j$  for some integers  $\kappa$  and  $\nu$  with  $\nu \geq \kappa$  can be written as  $\sigma^{\nu-\kappa} (\nu - \kappa)! \binom{\lambda/\sigma - \kappa}{\nu - \kappa}$ . But our assumption on  $\alpha$  means that  $c_l = \frac{c_0 h^l}{l!}$ , so that the two sums we have to compare amount to  $c_0^2 \sigma^{p-2k}$  times  $\sum_{n+m=p} \binom{\lambda/\sigma - k}{n-k} \binom{\lambda/\sigma - k}{m-k}$  and  $\sum_{n+m=p} \binom{\lambda/\sigma - 1 - k}{n-1-k} \binom{\lambda/\sigma + 1 - k}{m+1-k}$ . But after the appropriate translations of indices Lemma 5.2 shows that the two latter sums both equal  $\binom{2\lambda/\sigma - 2k}{p-2k}$ , so that  $A \in R_{\widetilde{W}}$  also under condition (ii).

Conversely, we first note that the valuation of both power series is  $2k$ , and that both have coefficient  $c_0^2$  in front of  $y^{2k}$  (as such terms are obtained only from  $n = m = k$  on the left hand side and from  $n = k + 1$  and  $m = k - 1$  on the right hand side). The coefficient of  $y^{2k+1}$  is  $2c_0 c_1 \gamma_k$  on the left hand side, while on the right hand side we get  $c_0 c_1 (\gamma_{k+1} + \gamma_{k-1})$ . Assuming that  $c_1 \neq 0$ , we find that the sequence  $\{\gamma_k\}_{k \in \mathbb{N}}$  satisfies the recursive relation which is satisfied by constant and linear sequences. Since the space of sequences satisfying a recursive relation of order 2 is 2-dimensional, we find that  $\gamma_k$  must be of the form  $\lambda - \sigma k$  for two constants  $\lambda$  and  $\sigma$  which are independent of  $k$ . Now, if  $\sigma = 0$  then  $\lambda \neq 0$  (since none of the  $\gamma_k$  may vanish), and we are in case (i). Assuming that  $\sigma \neq 0$ , it remains to prove that  $\mathbb{F}$  has characteristic 0 and that  $\alpha$  has the desired form. Let  $h = \frac{c_1}{c_0} \neq 0$ , and we write each  $c_l$  as  $c_0 h^l d_l$  for some  $d_l \in \mathbb{F}$ . We choose some  $k$ , and write the equation between the coefficients of  $y^p$  as follows: Take  $p = 2k + l$ , let  $\mu = \frac{\lambda}{\sigma} - k$ , and make the index change  $n = r + k$  and  $m = s + k$  on the left hand side while writing  $n = r + k + 1$  and  $m = s + k - 1$  on the right hand side. Substituting the value of  $\gamma_j$  and dividing the resulting equality by  $c_0^2 (\sigma h)^{p-2k=l}$  produces the equation

$$\sum_{r+s=l} d_r d_s \prod_{j=0}^{r-1} (\mu - j) \prod_{j=0}^{s-1} (\mu - j) = \sum_{r+s=l} d_r d_s \prod_{j=0}^{r-1} (\mu - 1 - j) \prod_{j=0}^{s-1} (\mu + 1 - j)$$

(which is independent of  $k$ ). The equalities for  $l = 0$  and  $l = 1$  are tautological, and  $d_0 = d_1 = 1$ . We now use the equation for larger  $l$  to prove that  $d_l$  satisfies the equality  $l! d_l = 1$  in  $\mathbb{F}$ , so that  $l!$  is invertible in  $\mathbb{F}$  and  $d_l = \frac{1}{l!}$ . Indeed, the induction hypothesis allows us to write  $d_r d_s = \frac{1}{r! s!}$  unless  $r = l$  and  $s = 0$  or the other way around, with the characteristic of  $\mathbb{F}$  allowing this. Our equality

then becomes

$$2d_l \prod_{j=0}^{l-1} (\mu-j) + \sum_{r+s=l} \binom{\mu}{r} \binom{\mu}{s} = d_l \prod_{j=-1}^{l-2} (\mu-j) + d_l \prod_{j=1}^l (\mu-j) + \sum_{r+s=l} \binom{\mu+1}{r} \binom{\mu-1}{s}.$$

The difference  $\prod_{j=-1}^{l-2} (\mu-j) - \prod_{j=0}^{l-1} (\mu-j)$  is  $l \prod_{j=0}^{l-2} (\mu-j) = l! \binom{\mu}{l-1}$  (we already know that  $(l-1)!$  is invertible in  $\mathbb{F}$ ), and the difference between  $\prod_{j=0}^{p-1} (\mu-j)$  and  $\prod_{j=1}^l (\mu-j)$  similarly equals  $l \prod_{j=1}^l (\mu-j) = l! \binom{\mu-1}{l-1}$ . Moreover, Corollary 5.3 allows us to write  $\binom{\mu+1}{r}$  as  $\binom{\mu}{r} + \binom{\mu}{r-1}$  on the right hand side as well as  $\binom{\mu}{s} = \binom{\mu-1}{s} + \binom{\mu-1}{s-1}$  on the left hand side (note that Lemma 5.2 and Corollary 5.3 are valid not only in characteristic 0, but wherever the factorial of the lower index is invertible in  $\mathbb{F}$ ). After cancelations we obtain

$$d_l l! \binom{\mu-1}{l-1} + \sum_{r+s=l} \binom{\mu}{r} \binom{\mu-1}{s-1} = d_l l! \binom{\mu}{l-1} + \sum_{r+s=l} \binom{\mu}{r-1} \binom{\mu-1}{s}.$$

Recalling that the indices  $r$  and  $s$  are assumed not to vanish, Lemma 5.2 shows that the sum on the left hand side is  $\binom{2\mu-1}{l-1} - \binom{\mu-1}{l-1}$ , while the one on the right hand side equals  $\binom{2\mu-1}{l-1} - \binom{\mu}{l-1}$ . After cancelation and using Corollary 5.3 once more we establish the equality  $\binom{\mu-1}{l-2} = d_l l! \binom{\mu-1}{l-2}$ , from which the desired equality follows since the non-vanishing of all of the numbers  $\gamma_k$  implies that the latter extended binomial coefficient is non-zero. Hence  $\mathbb{F}$  is of characteristic 0,  $d_l = \frac{1}{l!}$  for all  $l$ ,  $c_l = c_0 \frac{h^l}{l!}$  for all  $l$ , and  $\alpha(y) = c_0 e^{hy}$  as required. This completes the proof of the theorem.  $\square$

**Corollary 5.5.** *Take  $W(t) = e^t$  (in characteristic 0). Then a  $W$ -Appell sequence with  $p_1(0) \neq 0$  which corresponds to the matrix  $A \in R_W \subseteq L$  is  $\widetilde{W}$ -Sheffer if and only if either  $\widetilde{W}(t) = e^{t/\lambda}$ , or  $\widetilde{W}(t) = (1 + \sigma t)^{\lambda/\sigma}$  for some non-zero  $\sigma \in \mathbb{F}$  and  $\lambda \in \mathbb{F} \setminus \mathbb{N}\sigma$  and  $A$  is a non-zero scalar multiple of the matrix from Proposition 4.2 which represents a  $W$ -translation  $T_{h,W}$  for some  $h \in \mathbb{F}$ .*

*Proof.* We just substitute  $W(t) = e^t$  in Theorem 5.4. The case  $\widetilde{W}(t) = e^{t/\lambda}$  is case (i) there (or Proposition 5.1). Otherwise the requirement that  $\alpha(y) = c_0 e^{hy}$  is precisely the condition from Proposition 4.2 for our  $W$  (up to a multiplicative scalar). Moreover, the condition  $\gamma_k = \lambda - \sigma k$  yields here  $\frac{1}{w_{k+1}} = \frac{\lambda - \sigma k}{w_k(k+1)}$ . This shows, by a simple induction, that  $\frac{1}{w_k}$  is the coefficient  $\sigma^k \binom{\lambda/\sigma}{k}$  of  $(1 + \sigma t)^{\lambda/\sigma}$ . This proves the corollary.  $\square$

In fact, Proposition 4.6 allows us to translate the (simpler) assertion of Corollary 5.5 to the general case considered in Theorem 5.4.

We remark that in case  $c_1 = 0$  the results of Theorem 5.4 and Corollary 5.5 are more complicated. To give a rough idea of this, let  $d = v(\alpha(y) - \alpha(0))$  by the minimal index  $k > 0$  such that  $c_k \neq 0$ . Then a more detailed analysis of the proof of Theorem 5.4 shows that this linear condition holds not for the numbers  $\gamma_k$  but for  $\prod_{i=0}^{d-1} \gamma_{k+i}$ . This allows us to obtain relations only between

$\tilde{w}_k$  and  $\tilde{w}_{k+d}$  rather than between  $\tilde{w}_k$  and  $\tilde{w}_{k+1}$ . Indeed, note that for  $d = \infty$  and  $\alpha(y) = c_0$  we get just a scalar matrix, which represents a scalar multiple of the monomial sequence. As this sequence is Appell (hence Sheffer) with respect to every  $W$ , this illustrates the fact that as  $d$  grows, less restrictions on  $\widetilde{W}$  must be imposed for a  $W$ -Appell sequence to be  $\widetilde{W}$ -Sheffer.

## References

- [LM] Luzón, A., Morón, M, RECURRENCE RELATIONS FOR POLYNOMIAL SEQUENCES VIA RIORDAN MATRICES, Linear Algebra and its Applications, vol 433, 1422–1446 (2010).
- [R] Roman, S., THE UMBRAL CALCULUS, Pure and Applied Mathematics, vol 111, Academic Press, x+193pp (1984).
- [S] Sprugnoli, R., RIORDAN ARRAYS AND COMBINATORIAL SUMS, Disc. Math., vol 132 issue 1–3, 267–290 (1992).
- [SGWW] Shapiro, L. W., Getu, S., Woan, W. J., Woodson, L. C., THE RIORDAN GROUP, Disc. App. Math., vol 34, 229–239 (1991).
- [WW] Wang, T, Wang, W., GENERALIZED RIORDAN ARRAYS, Disc. Math., vol 308 issue 24, 6466–6500 (2008).

FACHBEREICH MATHEMATIK, AG 5, TECHNISCHE UNIVERSITÄT DARMSTADT,  
SCHLOSSGARTENSTRASSE 7, D-64289, DARMSTADT, GERMANY  
E-mail address: zemel@mathematik.tu-darmstadt.de